

## PERSPECTIVE



WILEY

# Event parsing and the origins of grammar

Klaus Zuberbühler<sup>1,2</sup>

<sup>1</sup>School of Psychology and Neuroscience,  
University of St Andrews, St Andrews,  
Scotland

<sup>2</sup>Institute of Biology, University of  
Neuchâtel, Neuchâtel

## Correspondence

Klaus Zuberbühler, Institute of Biology,  
University of Neuchâtel, Neuchâtel,  
Switzerland.

Email: klaus.zuberbuehler@unine.ch

## Funding information

Swiss National Science Foundation,  
Grant/Award Numbers: 310030\_185324,  
51NF40\_180888

**Edited by:** Juan-Carlos Gómez, Editor

## Abstract

How did grammar evolve? Perhaps a better way to ask the question is what kind of cognition is needed to enable grammar. The present analysis departs from the observation that linguistic communication is structured in terms of agents and patients, a reflection of how humans see the world. One way to explore the origins of cognitive skills in humans is to compare them with primates. A first approach has been to teach great apes linguistic systems to study their production in subsequent conversations. This literature has revealed considerable semantic competences in great apes, but no evidence for a corresponding grammatical ability, at least in production. No ape has ever created a sentence with an underlying causal structure of agency and patienthood. A second approach has been to study natural communication in primates and other animals. Here, there is intermittent evidence of compositionality, for example, a capacity to perform operations on semantic units, but again no evidence for an ability to refer to the causal structure of events. Future research will have to decide whether primates and other animals are simply unable to see the world as casually structured the way humans do, or whether they are just unable to communicate causal structures to others.

This article is categorized under:

Cognitive Biology > Evolutionary Roots of Cognition

Computer Science and Robotics > Artificial Intelligence

Linguistics > Evolution of Language

## KEYWORDS

animal cognition, meaning, primate communication, syntax, thematic roles

## 1 | EVOLUTION OF LANGUAGE

Human language is likely to have some of its evolutionary roots in a species ancestral to modern apes and humans. As linguistic competence is the product of different component skills a recurrent question is which ones date back to a common ancestor and which ones have evolved in the more recent human lineage (Fitch & Zuberbühler, 2013).

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Author. *WIREs Cognitive Science* published by Wiley Periodicals LLC.

Here, the focus is on the capacity for grammar, which is essential for the composition of meaning. Grammar is so natural to humans that it is difficult to imagine life without it. Young children develop it spontaneously, without instruction, and sometimes even in exceptional circumstances without auditory input (Senghas et al., 2004), so what are the evolutionary roots of this capacity?

One hypothesis is that grammatical competence is the result of the way natural events are perceived (Zuberbühler, 2020). To humans, events are not holistic percepts, but they have an underlying invisible structure of responsibility and causation. Things do not just happen, they are caused by agents that act upon patients and themes with mechanical or psychological force. This way of seeing the world is so powerful that it has been argued that humans have something akin to an evolved 'causality detector' (Kummer, 1995), whose outputs are then fed into communication. Languages differ in their grammars, in how they represent the basic causal structure of events as perceived by their speakers, such as by word order, declination, or affixation. In doing so, the agents usually obtain the main attention, but languages have further grammatical devices to shift the attention to patients, for example, by passive constructions. In "I have been stung by a wasp" the wasp remains the agent but no longer has subject (attention) status. Thus, grammar is external a mere consequence of how humans see the world, but where are the phylogenetic roots of this propensity? Evolution is a gradual process so operationally it is often illuminating to investigate primates, both under laboratory conditions and in the wild.

## 2 | LABORATORY STUDIES

### 2.1 | Human languages

The most direct way to compare species in their cognition is to teach them a common language and compare their resulting performance. In one early study, a chimpanzee named Viki was taught spoken English but, despite intense instruction, the chimp never produced anything important, apart from a few sound patterns with some resemblance to English words (e.g., mama, papa, cup; Hayes, 1952). Similar findings were also reported from other laboratories (Ladygina-Kohts, 1935 [2002]), which led to the conclusion that great apes (and possibly all other non-human primates) simply lacked the necessary motor control of the vocal apparatus to produce human speech. This general conclusion has recently been refined due to observations of considerable control of facial muscles (Lameira et al., 2014, 2017) but the fact remains that primates struggle with coordinating laryngeal and supralaryngeal activity.

Animal language studies were more successful once they switched to gestural signals, such as American Sign Language (ASL). Important contributions were by research with the female chimpanzee Washoe (Gardner et al., 1989). Similar to Viki, Washoe was raised like a human child, that is, allowed to form social relations with her caretakers and participating in joint activities. Washoe learned over 100 ASL signs and, according to the study directors, operant conditioning did not work well as a training method; Washoe learned better if the signs were not rewarded. Another unexpected finding was that Loulis (Washoe's adopted son) learned ASL signs without human instruction, apparently directly by observing Washoe (Gardner et al., 1989; Gardner & Gardner, 1969).

Although the evidence is somewhat anecdotal, Washoe used ASL signs spontaneously and appropriately, for example, she produced the sign DOG for any dog (Gardner & Gardner, 1984), suggesting that she had some understanding of the conceptual implications linked to the sign. More controversially, she occasionally produced (untaught) novel combinations. A well-cited example is Washoe signing "water" and "bird" upon seeing a swan, or "metal," "cup," "drink" for a thermos bottle. Although these observations were interesting, it is impossible to rule out that the ape simply referred to different facets of the same object rather than creating a new name. Apart from a lack of systematic data, project Washoe was haunted by accusations of Clever Hans effects, that is, that Washoe was responding inadvertently to experimenter cues, rather than engaging in proper conversations.

In a more systematic study, Project Nim, an infant chimpanzee was taught a similar number of gestural signs. Although Nim routinely produced sequences of signs, there was no evidence for any sort of grammar or sentence formation in the output (Terrace et al., 1979). In a typical situation, a caretaker would for instance sign "Do you want to eat an apple?" to which Nim responded with "Eat apple" or "Apple eat eat apple eat apple hurry apple hurry hurry" (Tomasello, 1994). Thus the general conclusion was that the ape did not understand what he was signing, other than that some signs led to rewards, similar to pigeons or rats in operant conditioning paradigms (Terrace et al., 1979). What seemed like dialogues was the ape imitating a caretaker and producing signs that were somewhat relevant to the situation (Tomasello, 1994). Another important observation in Nim and other language-trained animals was that signal production was mainly of an imperative nature. The apes

communicated mainly to request food or activities, but rarely ever to describe their own reality in declarative ways (Terrace et al., 1979). As a result, it was difficult to decide whether there was much understanding of categorial properties or thematic roles the signs were part of. The way that Nim was raised and taught may partly explain his failure to use signs in intelligent ways. Human children do not learn languages to obtain rewards but to get things done socially, and there is no reason to believe that this should be fundamentally different in animals.

## 2.2 | Artificial languages

A second approach was to teach animals artificial, symbol-based languages. Von Glasersfeld's lexigram system, Yerkish, was pioneering by consisting of over 100 abstract symbols that could be operated from a keyboard (Rumbaugh, Gill, & von Glasersfeld, 1973). A female chimpanzee called Lana was successfully trained with this system (Rumbaugh, von Glasersfeld, et al., 1973), which would allow her to interact with her caretakers and, anecdotally, this led to proper dialogues (Gill, 1977). In a more systematic study, however, Lana failed to correctly respond to simple noun categories (food, tools), unable to sort novel food and non-food items into the corresponding bins, suggesting that she used lexigrams without comprehending their representational features (Savage-Rumbaugh et al., 1980). This suggested that the lexigram production of Lana was no different from sequences of key-pecks produced by an operantly conditioned laboratory pigeon (Terrace et al., 1977).

More remarkable was the performance of a bonobo, Kanzi. His mother Matata was exposed to years of lexigram training, albeit without much success. Kanzi was present during this training and apparently learned the lexigrams spontaneously, through mere observation learning and without formal training and reinforcement (Savage-Rumbaugh & Lewin, 1994). A second astonishing outcome was that Kanzi began to understand English, much beyond single word utterances and involving entire sentences (Savage-Rumbaugh et al., 1993). In systematic tests, Kanzi was exposed to English sentences in which the same words were presented in different ways, such as

A "Could you take the pine needles outdoors?" vs. "Go outdoors and get the pine needles!" In this type of comparison both word order and verb changed. Kanzi performed well with such requests, suggesting he understood the meaning of the sentences.

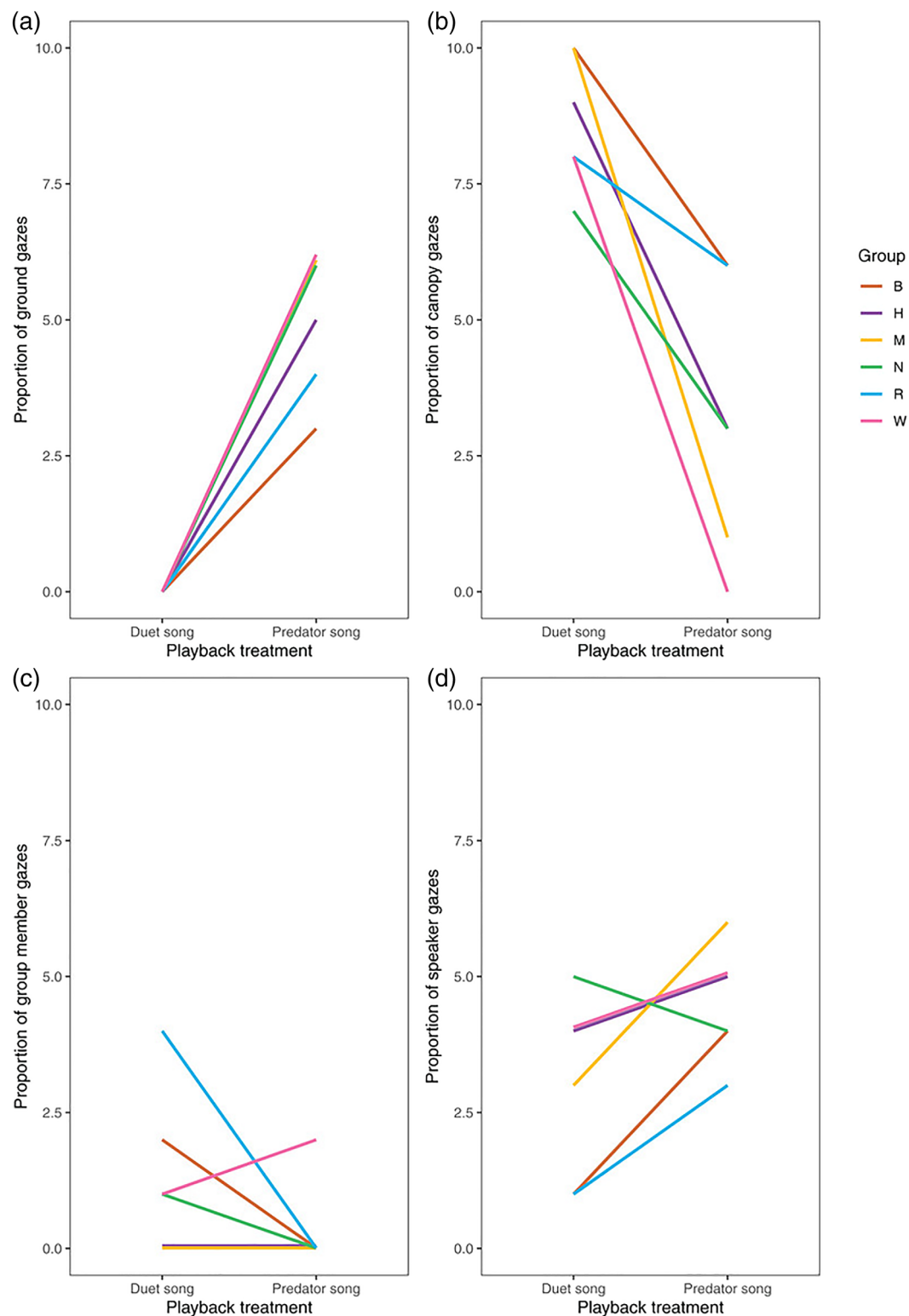
B "Take the rock outdoors!" vs. "Go get the rock that's outdoors!" Here, "take" and "get" mean different things, but "rock" has the same syntactic position (i.e., same word order) as the direct object. Kanzi was able to respond above chance, suggesting he understood not only the meaning of the verbs but also their thematic roles.

C "Make the doggie bite the snake" vs. "Make the snake bite the doggie." Here, the word order changes but the verb remains identical. Again Kanzi's performance was above chance in responding correctly, suggesting he was sensitive to word order. He appears to know that the actor is mentioned before the action while the patient is mentioned after.

Whether or not Kanzi really comprehended his lexigrams as organized in grammatical classes, such as nouns, verbs, agents, or patients, is an intriguing possibility that should be addressed by further research. Tomasello (1994) proposed to let Kanzi observe a familiar agent performing an *unfamiliar action* to a familiar patient, while hearing a name for the novel action ("Look! *Daxing!*"). If Kanzi had a notion of agent/patienthood he should know the correct order for mentioning the two participants in relation to "daxing." This is difficult for young children, who tend to randomly position agents and patients when producing novel verbs (Olguin & Tomasello, 1993), suggesting it might be beyond the cognitive capacities of great apes. The question of whether Kanzi, or any other great ape, is capable of extracting abstract sentence patterns across events, that is, whether he is able to form grammatical categories of agents, patients, or verbs is still unresolved.

## 2.3 | Artificial grammars

A third approach to study grammatical skills in animals has been with artificial grammar experiments (Fitch & Friederici, 2012). They differ in fundamental ways from the previous animal language studies insofar as the stimulus sequences are deliberately stripped off all semantic content. The idea is that this will allow researchers to study syntax in its pure,



**FIGURE 1** Responses of wild white-handed gibbon groups to playbacks of conspecific duet songs and predator songs (reprinted from Andrieu et al., 2020)

uncontaminated form, that is, the cognition of syntax. In a typical experiment, subjects are trained on syllable sequences organized along different structural complexities. For example, syllables can either be presented in a linear (finite state) way, for example, *pa-ba; li-di; mo-yo* or in a hierarchical (phrase structure) way, for example, *pa-li-mo; ba-di-yo* (Figure 1). Syntactic competence is then assessed testing whether subjects perceive violations of the rule they were taught before. The general finding from this research is that humans outperform animals with hierarchically organized grammars, which has

been taken to suggest that grammar is only possible after reaching a threshold in computational power (Fitch & Hauser, 2004). Although this line of research is important, it has the drawback that, in nature, syntax always operates alongside semantics in inseparable ways, raising questions about the relevance of this research for evolutionary questions.

### 3 | NATURAL COMMUNICATION

Kanzi's level of understanding English has provided support to the hypothesis that great apes have some capacity for grammar, a gradual cognitive achievement that is somewhat less pronounced compared to humans (Kako, 1999). If correct, we might expect similar capacities in the natural communication of bonobos and other great apes. However, Tomasello (1994) promoted a more pessimistic outlook, arguing that Kanzi's performance was the result of being raised in a human-like cultural environment from a young age. Apes in the wild, he argues, are unable to construct the necessary cultural or instructional environments on their own and will therefore not create the environment necessary for the emergence of linguistic communication. Grammar, in this view, is a linguistic competence that can only emerge in a human-like cultural environment. The following sections examine this claim, by reviewing some of the fieldwork conducted on natural primate communication systems.

#### 3.1 | Animal song

The notion of syntax has first been raised in studies of birdsong, for example, Bremond (1968). Here, syntax is typically defined in broad terms, such as rules that determine the order of elements. The research goal is to describe these rules and to see what information is encoded at different levels of organization (Kershenbaum et al., 2012).

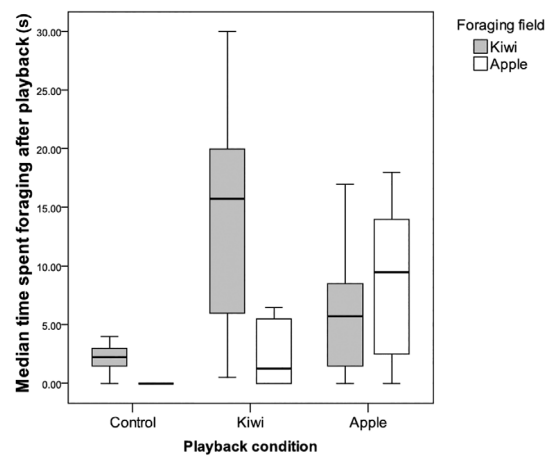
Animal song typically functions in reproduction, either to attract mating partners or to repel sexual rivals, a sexually selected behavior (Anderson, 1994). The songs of humpback whales are an intriguing example, with males combining units into phrases and songs, aided by social learning within groups and sometimes even between populations (Garland et al., 2017). In some bat species, such as Brazilian free-tailed bats, males also produce songs, which can be region-specific, but differ individually in syllable and phrase number (Bohn et al., 2009). Although such examples are enlightening, these syntaxes do not function to convey meaning, other than to promote the caller himself.

In primates, singing is uncommon but it has been reported in socially monogamous species, such as gibbons or indris. White-handed gibbons are well studied, because they use songs not only in social interactions but also as alarms to predators, notably tigers, clouded leopards, and pythons (Clarke et al., 2006). Predator and duet songs differ in their combinations but are assembled from the same basic song units, which can be discriminated by listeners (Andrieu et al., 2020). Although intriguing, individual song units have no specific meaning. Hence, although gibbon song conveys information about events external to the caller, it does not qualify as syntax in a linguistic sense.

#### 3.2 | Call combinations

There are multiple examples in animal communication of meaningful utterances composed of smaller meaningful elements, either in the form of unordered combinations or ordered permutations. One example is titi monkey alarm calling with adults producing A and B calls to aerial and terrestrial predators, respectively (Caesar et al., 2012). Furthermore, callers produced mixed sequences if the two predator classes were detected in non-standard locations, for example, a raptor on the ground or terrestrial predator in the canopy. For example, positioning an ocelot model into the canopy triggered series of B-calls that preceded by an A call (Cäsar et al., 2013). Although the sequences did not have clear permutation rules, they elicited adequate behavioral responses in playback experiments (Berthet et al., 2019; Cäsar et al., 2012). Most likely, listeners reacted to the proportion of B-call pairs (bigrams) relative to all call pairs of the sequence, suggesting that bigrams served as main carriers of meaning (Berthet et al., 2019).

Combinatorial syntax has also been found in great apes. Bonobos produce five call types to food (bark, B; peep, P; peep-yelp, PY; yelp, Y; grunt, G) as part of seemingly unordered call combinations (Clay & Zuberbühler, 2009). As with the titi monkeys, individual calls carry meaning, insofar as different calls are preferably given to some foods but not others. In a playback experiment, subjects first learned to find highly preferred food (kiwi) in one location of their enclosure and less preferred food (apple) in another location. They then heard playbacks with segments of call



**FIGURE 2** Responses of bonobos to playbacks of conspecific food call sequences to kiwis and apples. Reproduced with permission from Clay and Zuberbühler (2011) under the Creative Commons Attribution license

**TABLE 1** Summary information of food type used to elicit call sequences and subsequent receiver responses to the sequences (reprinted from Zuberbühler, 2020)

Bigram	Kiwi search (s) (median)	Apple search (s) (median)	Kiwi bias (median)	Apple bias (median)	N
B	20.5	3.8	5.1	1.3	4
P	20.8	6.5	5.2	1.2	3
PY	5.8	9.0	1.5	2.9	5 <sup>a</sup>
Y	5.2	12.3	1.3	4.2	4 <sup>a</sup>
n.a.	9.3	2.0	5.7	1.2	1
n.a.	6.5	17.3	1.4	3.7	1

<sup>a</sup> 1 of 15 sequence contained both Y and PY bigrams; n.a., no bigrams within the first four calls.

sequences originally given by others to either kiwi or apple. Although sequences consisted of only four calls, subjects responded in highly structured ways: sequences originally given to kiwi triggered mainly searching in the kiwi field whereas the opposite was observed with sequences originally given to apples (Figure 2) (Clay & Zuberbühler, 2011).

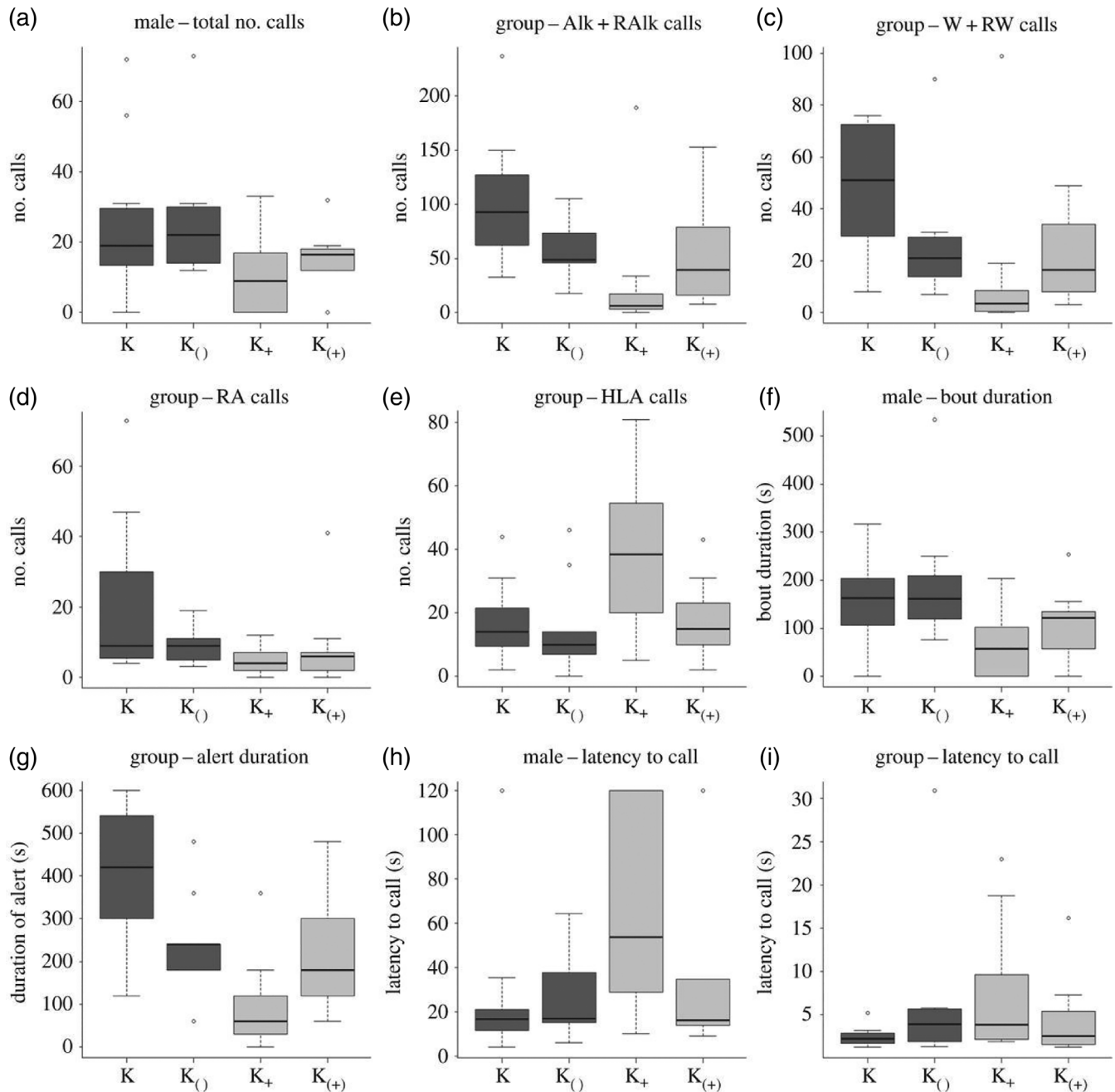
Are bonobo food calls compositional, that is, is the meaning of the utterance determined by the meaning of the constituents and the rule that combines them? Or, conversely, do words obtain their meanings from the meaning of the sentence in which they occur as constituents, Frege's context principle (Frege, 1884 [1950]), section 60. The hypothesis put forward by Clay and Zuberbühler (2011) was that that combined meaning of the bonobo food call sequences is not compositional but an amalgam of meanings of individual calls. In this view, listeners somehow integrate individuals meanings in an additive fashion before deciding whether a caller is responding to high or low quality food.

In titi monkeys, the proportion of B-call bigrams explained much of the monkeys' production and response patterns (Berthet et al., 2019) and it is possible that in bonobos, meaning is determined by the same principle. Post hoc analysis of the call sequences used as playback stimuli revealed that 15 of 17 of them contained bigrams of B, P, PY, and Y calls (Zuberbühler, 2020) (Table 1). In particular, kiwi consistently elicited call sequences that contained either B or P bigrams (6 of 7 sequences) while apple triggered sequences with PY and Y bigrams (8 of 10 sequences) within the first 4 calls. Also, listener responses were better explained by the distribution of bigrams than the food type that originally triggered the calls: Two sequences contained no bigrams and one contained contradictory information (B and PY bigrams). For the remaining unambiguous sequences, 12 of 14 (85.7%) triggered foraging behavior in the predicted way, whereas 12 of 17 (70.6%) triggered foraging behavior in the same food field as the food the calls were originally given to.

Why should callers produce bigrams rather than simply allocating single call types to single food types? One possibility is that the function of some calls is more general. This is certainly the case for P (peep) calls, which are also given to non-food related events and seem to function as a sort of vocal attention getter (Clay et al., 2015). Producing bigrams



allows callers to generate structures that are acoustically more conspicuous than single calls. As already argued for chimpanzees (Slocombe & Zuberbühler, 2006), bonobos feed on hundreds of items in the wild, suggesting complex call systems may have evolved to allow callers to convey variables beyond food desirability. Encoding individual identity, for example, could be of high importance, especially during foraging when reciprocity becomes important. With a simple duplication rule callers could generate extra meaning, beyond what is already available from their vocal repertoire. Repetitions of single calls also have semantic importance in black-and-white colobus alarm calls (Schel et al., 2009), suggesting that such operations may be of wider relevance in primate communication. Further work is needed to explore these questions.



**FIGURE 3** Median and inter-quartile range in the four experimental conditions natural 'Krak' (K, N = 11), artificial 'Krak' (K<sub>(-)</sub>, N = 9), natural 'Krak-oo' (K<sub>(+)</sub>, N = 12) and artificial 'Krak-oo' (K<sub>(+-)</sub>, N = 10) for the number of calls given by male Diana monkeys

### 3.3 | Call permutations

Evidence for permutations has been found in a number of primate communication studies, both in terms of within-call organization and at the sequence level. In Diana monkey contact calls, for example, the arched unit conveys the identity of the caller, which is more or less pronounced depending on the visibility and surrounding noise, suggesting callers have some control over articulation (Coye et al., 2016, 2018). Arches can have trill or chatter prefixes, depending on whether social interactions are positive (trills) or negative (chatter). Arches, trills, and chatters can either be produced individually or as merged permutations, but it is currently unknown whether the meaning of the combined calls is additive or compositional.

Another affixation example is Campbell's monkey alarm calls. Here, males produce three alarm call types in situations of danger, "krak," "hok," "wak." Each of the calls can be given either on its own or suffixed to "krak-oo," "hok-oo," or "wak-oo" (Ouattara et al., 2009b). Suffixation is not additive but compositional, insofar as suffixed calls are given to distinct situations different from unsuffixed ones: While "kraks" are given to leopards, "krak-oo"s are given to a wide range of unspecific disturbances, such as falling trees or other monkeys' distant alarm calls. Overall, suffixed calls are associated with low urgency situations, while unsuffixed calls are given when danger is imminent. In a playback study, krak alarms caused leopard alarm calls in associated groups of Diana monkeys whereas krak-oo alarms did not, regardless of whether the stimuli were naturally or artificially suffixed (Coye et al., 2015) (Figure 3). As the -oo affix is never produced on its own, but causes a semantic operation, it functions as a morpheme.

When responding to external events, Campbell's monkey males usually produce long sequences of alarm calls, as mentioned before. In addition, they can produce low-frequency boom calls prior to these sequences, but only to non-predatory disturbances, such as sudden loud noises (Ouattara et al., 2009a; Zuberbühler, 2001). In playback experiments, Campbell's monkey alarm call sequences were played back to groups of Diana monkeys, which responded by producing their own corresponding alarm calls. However, if boom calls were added before the alarm call series, Diana monkeys ceased to alarm call, suggesting that the booms had modified the meaning of the entire call sequence, rendering them from specific predator alarms to unspecific disturbance calls (Zuberbühler, 2002).

Semanticity has also been found in putty-nosed monkey alarm calls. Here, males produce sequences of pyows to general disturbances and sequences of hacks to crowned eagles. In addition they can give mixed sequences of pyows followed by hacks prior to group movement and hacks followed by pyows after eagle encounters (Arnold & Zuberbühler, 2006). In playback experiments, pyows triggered gazes towards the presumed caller, suggesting listeners searched for additional information, whereas hacks triggered gazes towards the upper canopy, suggesting listeners were looking for an eagle. Pyow-hack sequences, finally, triggered movements towards the speaker (i.e., the presumed caller) (Arnold & Zuberbühler, 2008).

Does the putty-nosed monkey call system qualify as compositional? One hypothesis proposes that pyow-hack sequences are syntactically combinatorial but not semantically compositional because their meaning cannot be derived from the meanings of the component calls (Arnold & Zuberbühler, 2012). Pyow-hack sequences, in this view, function similar to idioms, by conveying meaning that is not directly extractable from the constituent parts (e.g., *up in the air* = unresolved).

Another hypothesis allows for a compositional interpretation, based on the premise that calls are produced along an urgency principle (Schlenker et al., 2016). The principle states that calls with information about the nature/location of a threat (i.e., hacks) are more urgent and come before calls that do not (i.e., pyows). In this sense, hack series are urgent instructions (e.g., to avoid the eagle), whereas pyow series are non-urgent instructions (e.g., to locate the caller). Logically, then, *hack-pyow* sequences are transitions from an urgent to a non-urgent state, whereas *pyow-hack* sequences can only indicate a non-urgent state, as the non-urgent call is produced first. Semantically, pyow-hack sequences are compatible with any kind of arboreal movement.

A third and hitherto unexplored hypothesis is that pyows function as general alerts and attention getters which, similar to the Campbell's monkey booms, remove the referential specificity of subsequent (hack) alarm calls by rendering them from specific eagle to generalized aerial danger alarms. Males announce forthcoming group movements this way because traveling through the canopy increases the likelihood of encountering a crowned eagle. Pyow-hack combinations, in other words, refer to the *possibility* rather than the *reality* of encountering an eagle. Similar observations have been made in titi monkeys who produce B-alarm calls when detecting a ground predator or when descending towards the ground (Berthet et al., 2018). According to this theory, males produce pyow-introduced hack series to distinguish likely from real predator encounters.

Currently, it is not possible to discriminate between the three hypotheses. One way to put them to a test would be to study putty-nosed monkey populations in habitats with high leopard predation. For example, the urgency principle predicts that in leopard-infested populations "pyows" (or a version thereof) will convey specific information about



location/nature of the leopard threat, similar to how “hacks” refer to eagle threats. Pyows vary in their acoustic structure, suggesting that callers in leopard-infested habitats could benefit from this variation to produce high-urgency, leopard-related pyows and also low-urgency (semantic operator) pyows to announce group movement. But even without acoustic variation pyows could still function either as high urgency leopard alarms (if produced alone) or as semantic modifiers for rendering subsequent hacks from high-urgency eagle alarms to low-urgency (anticipated) canopy threats. More research is needed to address the different possibilities.

For great apes, Crockford and Boesch (2005) have found numerous call combinations that could be associated with some emission contexts but not others. Regarding permutations, the currently best example is probably the pant-hoot calls of chimpanzees, a vocal utterance produced in long-distance communication. Pant hoots are composed of four discrete units, the introduction, build-up, climax and let-down, in this order. Using support vector machine methodology it was found that caller identity was apparent in all four phases, but most strongly in the low-amplitude introduction and high-amplitude climax phases. The age of the callers was also apparent in the introduction and build-up phases, whereas dominance rank was mainly apparent in the climax phase, and the activity of the caller mainly in the let-down phase (Fedurek et al., 2016). Future research should investigate whether the four call units are given independently and in what context to decide whether pant-hoots qualify as truly compositional signals.

## 4 | CONCLUSIONS

Can studies of primate communication advance our understanding of the evolution of human grammar? Regarding this question, Tomasello (1994) is unusually candid: “Linguists and other scientists will learn much more about processes of linguistic communication by studying apes such as Kanzi than they will by studying computer programs written by colleagues” (p. 388). Nevertheless, the studies reviewed here show that, even under favorable circumstances, great apes and other primates are unable to construct something akin to a sentence. Even human-enculturated apes use acquired communication systems mainly as imperative tools, to lodge request to their caretakers. They do this in seemingly unordered and unstructured ways, without much evidence of a capacity for grammar. In particular, there is no evidence that apes see any necessity in communicating agent/patient relationships, the foundation of all human languages. Kanzi is somewhat of an exception because of his demonstrated capacity to understand agent/patienthood from English sentences, although there is no corresponding evidence for production.

Clearly, sign languages and lexigram systems are alien to primates (and humans) and not the default way by which either species would interact naturally, but it is still striking how little creative use great apes have made and how little some of them appeared to understand these systems. Yet, the sample size is small and it is possible that some individuals were simply less apt than others. In some cases, particularly Nim, the teaching methods and social environment possibly interfered and prevented a fuller realization of the subject’s intellectual capacity.

When looking at natural communication the evidence is more promising, with multiple studies showing elements of compositionality. Some evidence is in terms of permutations, other evidence in terms of duplications (bigrams) of single calls within seemingly unordered larger sequences. However, no animal study so far has produced evidence that subjects communicate to refer to entire events, with agents and patients, the way humans do. Instead, animal signal compositions are shorthand references, either to the self (e.g., to make the caller’s behavior more predictable) or to important events (e.g., discovery of food or a predator) but with no evidence of conveying anything about the causal structure of the perceived event. Either this is because primates are unable to perceive causality in natural events or because they are not motivated to communicate it.

The conclusion is that the origins of grammar lie in the uniquely human cognition of assigning causality to any natural event, by which agents are responsible for allocating effects on patients. There is no evidence for such processes in animal communication, but there is evidence for simple semantic operations by which the meaning of larger compounds can be derived from the meaning of constituent parts, suggesting that this type linguistic capacity is phylogenetically older than human. Future research on the origins of grammar should focus on how animals perceive natural events and whether they are capable of communicating their analyses to others.

## ACKNOWLEDGMENTS

I am grateful to the Swiss National Science Foundation for funding this research (Project grant 310030\_185324 and NCCR Evolving Language, 51NF40\_180888) and to Balhasar Bickel and Philippe Schlenker for discussions.

Open access funding provided by Universite de Neuchatel.

Open access funding enabled and organized by Projekt DEAL.

## CONFLICT OF INTEREST

The author has declared no conflicts of interest for this article.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## ORCID

Klaus Zuberbühler  <https://orcid.org/0000-0001-8378-088X>

## RELATED WIREs ARTICLE

[Linguistic capacity of non-human animals](#)

## REFERENCES

- Anderson, M. (1994). *Sexual selection*. Princeton University Press.
- Andrieu, J., Penny, S. G., Bouchet, H., Malaivijitnond, S., Reichard, U. H., & Zuberbühler, K. (2020, August). White-handed gibbons discriminate context-specific song compositions. *PeerJ*, 8, e9477. <https://doi.org/10.7717/peerj.9477>
- Arnold, K., & Zuberbühler, K. (2006). Semantic combinations in primate calls. *Nature*, 441(7091), 303–303.
- Arnold, K., & Zuberbühler, K. (2008, March 11). Meaningful call combinations in a non-human primate. *Current Biology*, 18(5), R202–R203. <https://doi.org/10.1016/j.cub.2008.01.040>
- Arnold, K., & Zuberbühler, K. (2012, March). Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language*, 120(3), 303–309. <https://doi.org/10.1016/j.bandl.2011.10.001>
- Berthet, M., Mesbahi, G., Pajot, A., Cäsar, C., Neumann, C., & Zuberbühler, K. (2019). Titi monkeys combine alarm calls to create probabilistic meaning. *Science Advances*, 5, eaav3991. <https://doi.org/10.1126/sciadv.aav3991>
- Berthet, M., Neumann, C., Mesbahi, G., Casar, C., & Zuberbühler, K. (2018, January). Contextual encoding in titi monkey alarm call sequences. *Behavioral Ecology and Sociobiology*, 72(1), 8. <https://doi.org/10.1007/s00265-017-2424-z>
- Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M., & Pollak, G. D. (2009). Versatility and stereotypy of free-tailed bat songs. *PLoS One*, 4, e6746. <https://doi.org/10.1371/journal.pone.0006746>
- Bremond, J. C. (1968). Specific value of syntax in territorial defense signal of wren (*Troglodytes troglodytes*). *Behaviour*, 30, 66–75. <https://doi.org/10.1163/156853968x00180>
- Caesar, C., Byrne, R., Young, R. J., & Zuberbühler, K. (2012, May). The alarm call system of wild black-fronted titi monkeys, *Callicebus nigrifrons*. *Behavioral Ecology and Sociobiology*, 66(5), 653–667. <https://doi.org/10.1007/s00265-011-1313-0>
- Cäsar, C., Byrne, R. W., Hoppitt, W., Young, R. J., & Zuberbühler, K. (2012, August). Evidence for semantic communication in titi monkey alarm calls. *Animal Behaviour*, 84(2), 405–411. <https://doi.org/10.1016/j.anbehav.2012.05.010>
- Cäsar, C., Zuberbühler, K., Young, R. J., & Byrne, R. W. (2013, October 23). Titi monkey call sequences vary with predator location and type. *Biology Letters*, 9(5), 20130535. <https://doi.org/10.1098/rsbl.2013.0535>
- Clarke, E., Reichard, U., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS One*, 1(1), e73.
- Clay, Z., Archbold, J., & Zuberbühler, K. (2015, August 4). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, 3, e1124. <https://doi.org/10.7717/peerj.1124>
- Clay, Z., & Zuberbühler, K. (2009, June). Food-associated calling sequences in bonobos. *Animal Behaviour*, 77(6), 1387–1396. <https://doi.org/10.1016/J.Anbehav.2009.02.016>
- Clay, Z., & Zuberbühler, K. (2011, April 27). Bonobos extract meaning from call sequences. *PLoS One*, 6(4), e18786. <https://doi.org/10.1371/journal.pone.0018786>
- Coye, C., Ouattara, K., Arlet, M. E., Lemasson, A., & Zuberbühler, K. (2018, July). Flexible use of simple and combined calls in female Campbell's monkeys. *Animal Behaviour*, 141, 171–181. <https://doi.org/10.1016/j.anbehav.2018.05.01>
- Coye, C., Ouattara, K., Zuberbühler, K., & Lemasson, A. (2015, May). Suffixation influences receivers' behaviour in non-human primates. *Proceedings of the Royal Society B – Biological Sciences*, 282(1807), 20150265. <https://doi.org/10.1098/rspb.2015.0265>
- Coye, C., Zuberbühler, K., & Lemasson, A. (2016, May). Morphologically structured vocalizations in female Diana monkeys. *Animal Behaviour*, 115, 97–105. <https://doi.org/10.1016/j.anbehav.2016.03.010>
- Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour*, 142, 397–421.
- Fedurek, P., Zuberbühler, K., & Dahl, C. D. (2016). Sequential information in a great ape utterance. *Scientific Reports*, 6, 38226.
- Fitch, W. T., & Hauser, M. D. (2004, Jan). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303(5656), 377–380.
- Fitch, W. T., & Friederici, A. D. (2012, July). Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B – Biological Sciences*, 367(1598), 1933–1955. <https://doi.org/10.1098/rstb.2012.0103>
- Fitch, W. T., & Zuberbühler, K. (2013). Primate precursors to human language: Beyond discontinuity. In E. Altenmüller, S. Schmidt, & E. Zimmermann (Eds.), *Evolution of emotional communication: From sounds in nonhuman mammals to speech and music in man*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199583560.003.0002>
- Frege, G. (1884 [1950]). *The foundations of arithmetic* [Die Grundlagen der Arithmetik] (J. L. Austin, Trans.). Blackwell.

- Garland, E. C., Rendell, L., Lamoni, L., Poole, M. M., & Noad, M. J. (2017, Jul). Song hybridization events during revolutionary song change provide insights into cultural transmission in humpback whales. *Proceedings of the National Academy of Sciences of the United States of America*, 114(30), 7822–7829.
- Gardner, R., & Gardner, B. (1984). A vocabulary test for chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 98(4), 381–404.
- Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, 165, 664–672.
- Gardner, R. A., Gardner, B. T., & Van Cantfort, T. E. (Eds.). (1989). *Teaching sign language to chimpanzees*. State University of New York.
- Gill, T. V. (1977). Conversations with Lana. In D. M. Rumbaugh (Ed.), *Language learning by a chimpanzee, the LANA project* (pp. 225–246). Academic Press.
- Hayes, C. (1952). *The ape in our house*. Gollacz.
- Kako, E. (1999). Elements of syntax in the systems of three language-trained animals. *Animal Learning & Behavior*, 27(1), 1–14.
- Kershenbaum, A., Ilany, A., Blaustein, L., & Geffen, E. (2012, August 7). Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B – Biological Sciences*, 279(1740), 2974–2981. <https://doi.org/10.1098/rspb.2012.0322>
- Kummer, H. (1995). Causal knowledge in animals. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 26–39). Clarendon Press.
- Ladygina-Kohts, N. N. (1935 [2002]). *Infant chimpanzee and human child* (B. Vekker, Trans.; F.B.M. de Waal, Ed.). Oxford University Press.
- Lameira, A. R., Maddieson, I., & Zuberbühler, K. (2014, February). Primate feedstock for the evolution of consonants. *Trends in Cognitive Sciences*, 18(2), 60–62. <https://doi.org/10.1016/j.tics.2013.10.013>
- Lameira, A. R., Vicente, R., Alexandre, A., Campbell-Smith, G., Knott, C., Wich, S., & Hardus, M. E. (2017, February). Proto-consonants were information-dense via identical bioacoustic tags to proto-vowels. *Nature Human Behaviour*, 1(2), 0044. <https://doi.org/10.1038/s41562-016-0044>
- Olguin, R., & Tomasello, M. (1993). Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development*, 8, 245–272.
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009a, December 22). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51), 22026–22031.
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009b). Campbell's monkeys use affixation to alter call meaning. *PLoS One*, 4(11), e7808. <https://doi.org/10.1371/journal.pone.0007808>
- Rumbaugh, D. M., Gill, T. V., & von Glasersfeld, E. (1973). Reading and sentence completion by a chimpanzee (Pan). *Science*, 182, 731–733.
- Rumbaugh, D. M., von Glasersfeld, E., Warner, H., Pisani, P., Gill, T. V., Brown, J. V., & Bell, C. L. (1973). Exploring language skills of Lana chimpanzee. *International Journal of Symbolology*, 4(2), 1–9.
- Savage-Rumbaugh, E. S., Rumbaugh, D. M., Smith, S. T., & Lawson, J. (1980). Reference: The linguistic essential. *Science*, 210, 922–925.
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, 58(3–4), 1–222.
- Savage-Rumbaugh, S., & Lewin, R. (1994). *Kanzi*. Wiley.
- Schel, A. M., Tranquilli, S., & Zuberbühler, K. (2009, May). The alarm call system of two species of black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*). *Journal of Comparative Psychology*, 123(2), 136–150.
- Schlenker, P., Chemla, E., Arnold, K., & Zuberbühler, K. (2016, February). Pyow-hack revisited: Two analyses of Putty-nosed monkey alarm calls. *Lingua*, 171, 1–23. <https://doi.org/10.1016/j.lingua.2015.10.002>
- Senghas, A., Kita, S., & Ozyurek, A. (2004, September 17). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, 305(5691), 1779–1782.
- Slocombe, K. E., & Zuberbühler, K. (2006). Food-associated calls in chimpanzees: Responses to food types or food preferences? *Animal Behaviour*, 72, 989–999.
- Terrace, H. S., Petitto, L. A., Sanders, D. L., & Bever, T. G. (1979). Can an ape create a sentence? *Science*, 206, 891–902.
- Terrace, H. S., Straub, R., Bever, T. G., & Seidenberg, M. (1977). Representation of a sequence by pigeons. *Bulletin of the Psychonomic Society*, 10(4), 269–269.
- Tomasello, M. (1994, October). Can an ape understand a sentence – A review of language comprehension in ape and child. *Language & Communication*, 14(4), 377–390.
- Zuberbühler, K. (2001, October). Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus campbelli*. *Behavioral Ecology and Sociobiology*, 50(5), 414–422. <https://doi.org/10.1007/S002650100383>
- Zuberbühler, K. (2002, February). A syntactic rule in forest monkey communication. *Animal Behaviour*, 63, 293–299. <https://doi.org/10.1006/Anbe.2001.1914>
- Zuberbühler, K. (2020, January). Syntax and compositionality in animal communication. *Philosophical Transactions of the Royal Society B – Biological Sciences*, 375(1789), 20190062. <https://doi.org/10.1098/rstb.2019.0062>

**How to cite this article:** Zuberbühler, K. (2021). Event parsing and the origins of grammar. *Wiley Interdisciplinary Reviews: Cognitive Science*, e1587. <https://doi.org/10.1002/wcs.1587>